Proportional Size Distribution (PSD)

Alex J. Benecke

January 10, 2018

I will use a Chi-Square test to see if there is a difference between PSD between years (2013 - 2017).

Data Preparation

Load Data

```
lmbs <- read.csv("Data/Clean-Data/2012-2017_nearshore-survey-largemouth-bass_Stock_CLEAN.csv") %>%
    filter(Year >= 2013) %>% arrange(Year, FID, Length)
lmbs$fyr <- as.factor(lmbs$fyr)

str(lmbs)
```

```
## 'data.frame':    447 obs. of  16 variables:
##  $ Year   : int  2013 2013 2013 2013 2013 2013 2013 2013 2013 2013 ...
##  $ Site   : int  11 11 11 11 11 11 11 10 10 10 ...
##  $ FID    : int  1 2 3 4 6 7 8 9 10 17 ...
##  $ Weight: num  NA NA NA NA NA NA NA NA NA NA ...
##  $ Length: int  395 348 266 224 318 273 426 387 264 291 ...
##  $ AC     : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ AGE    : int  3 2 1 1 2 1 3 4 1 1 ...
##  $ SexCon: int  3 8 8 6 8 8 3 8 3 8 ...
##  $ Sex    : int  1 2 2 2 2 2 1 2 1 2 ...
##  $ Delts  : logi  NA NA NA NA NA NA ...
##  $ logW   : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ logL   : num  2.6 2.54 2.42 2.35 2.5 ...
##  $ fyr    : Factor w/ 5 levels "2013","2014",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ Ws     : num  935 617 256 146 460 ...
##  $ Wr     : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ gcat   : Factor w/ 3 levels "preferred","quality",..: 1 2 3 3 2 3 1 1 3 3 ...
```

```
headtail(lmbs)
```

```
##     Year Site FID Weight Length AC AGE SexCon Sex Delts     logW     logL
## 1   2013   11   1     NA    395 NA   3      3   1    NA       NA 2.596597
## 2   2013   11   2     NA    348 NA   2      8   2    NA       NA 2.541579
## 3   2013   11   3     NA    266 NA   1      8   2    NA       NA 2.424882
## 445 2017   12  NA   1362    435  3  NA     NA  NA    NA 3.134177 2.638489
## 446 2017   18  NA   1400    438  3  NA     NA  NA    NA 3.146128 2.641474
## 447 2017    9  NA   1362    464  3  NA     NA  NA    NA 3.134177 2.666518
##      fyr        Ws        Wr      gcat
## 1   2013  934.6786        NA preferred
## 2   2013  617.4316        NA   quality
## 3   2013  256.2345        NA     stock
## 445 2017 1281.6674 106.26782 preferred
## 446 2017 1310.8251 106.80296 preferred
## 447 2017 1583.1182  86.03274 preferred
```

```r
unique(lmbs$Year)  ### See that there is no 2012
```

```
## [1] 2013 2014 2015 2016 2017
```

View Data

```r
(lmbs.LF <- xtabs(~Year+gcat,data=lmbs))
```

```
##       gcat
## Year   preferred quality stock
##   2013        16      41    41
##   2014        18      57    65
##   2015        15      38    14
##   2016        11      44    52
##   2017         9      15    11
```

Chi-Squares Test

Is there a difference in the number of fish in each gabelhouse categorie during the years 2013 - 2017?

```r
chisq.test(lmbs.LF)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lmbs.LF
## X-squared = 20.055, df = 8, p-value = 0.01013
```

This seems to suggest that the proportional stock distribution (PSD) is different for largemouth bass between years ($X^2 = 20.055$, df = 8, P = 0.01013).

In which years is PSD different?

```r
round(prop.table(lmbs.LF,margin=1)*100,0)
```

```
##       gcat
## Year   preferred quality stock
##   2013        16      42    42
##   2014        13      41    46
##   2015        22      57    21
##   2016        10      41    49
##   2017        26      43    31
```

Remarkably the percent of quality fish is the same for 2014 and 2016 and a bit higher for 2015. the percentage of fish in each gcat is similar between years 2013, 2014, 2016, and 2017. However, the year 2015 appears to have a higher percentage of large fish and far fewer small fish.

1) Could this be some sort of sampling bias?
2) Could this be a result of sampling different sites? where the 2015 sites more suitable for LMB?
3) Are the years really different or do I just have too few to say for sure?

Compare PSD-Q between years 2013 - 2017

```
lmbs %<>% mutate(gcatQ=mapvalues(gcat,
                                 from=c("stock","quality","preferred"),
                                 to=c("quality-", "quality+", "quality+")),
                 gcatQ=droplevels(gcatQ))

(lmb.LFQ <- xtabs(~Year+gcatQ,data = lmbs))
```

```
##       gcatQ
## Year   quality+ quality-
##   2013       57       41
##   2014       75       65
##   2015       53       14
##   2016       55       52
##   2017       24       11
```

```
chisq.test(lmb.LFQ)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lmb.LFQ
## X-squared = 16.815, df = 4, p-value = 0.0021
```

```
(ps.Q <- c(chisq.test(lmb.LFQ[c(1,2),])$p.value,  ### 2013-2014
           chisq.test(lmb.LFQ[c(1,3),])$p.value,  ### 2013-2015
           chisq.test(lmb.LFQ[c(1,4),])$p.value,  ### 2013-2016
           chisq.test(lmb.LFQ[c(1,5),])$p.value,  ### 2013-2017
           chisq.test(lmb.LFQ[c(2,3),])$p.value,  ### 2014-2015
           chisq.test(lmb.LFQ[c(2,4),])$p.value,  ### 2014-2016
           chisq.test(lmb.LFQ[c(2,5),])$p.value,  ### 2014-2017
           chisq.test(lmb.LFQ[c(3,4),])$p.value,  ### 2015-2016
           chisq.test(lmb.LFQ[c(3,5),])$p.value,  ### 2015-2017
           chisq.test(lmb.LFQ[c(4,5),])$p.value)) ### 2016-2017

(p.val.Q <- p.adjust(ps.Q))
```

```
##  [1] "13-14" "13-15" "13-16" "13-17" "14-15" "14-16" "14-17" "15-16"
##  [9] "15-17" "16-17"
```

```
##     Year p-value Adjusted p
## 1  13-14  0.5694     1.0000
## 2  13-15  0.0084     0.0675
## 3  13-16  0.4060     1.0000
## 4  13-17  0.3781     1.0000
## 5  14-15  0.0007     0.0064
## 6  14-16  0.8338     1.0000
## 7  14-17  0.1583     0.9500
## 8  15-16  0.0005     0.0046
## 9  15-17  0.3515     1.0000
## 10 16-17  0.1144     0.8007
```

The PSD-Q of largemouth bass is different for at least one of the years during 2013 - 2016 (Chi-Squared, $X^2 = 16.815$, df = 4, p = 0.0021). The adjusted p-values show a *significant difference* in PSD-Q between years 2014 - 2015 (p = 0.0064) and 2015 - 2016 (p = 0.0046). The PSD-Q is not different between any other years. However 2013 and 2015 may be different (p = 0.0675).

Compare PSD-P between years 2013 - 2017

```
lmbs %<>% mutate(gcatP=mapvalues(gcat,
                                 from=c("stock","quality","preferred"),
                                 to=c("preferred-", "preferred-", "preferred+")),
                 gcatP=droplevels(gcatP))

(lmb.LFP <- xtabs(~Year+gcatP,data = lmbs))
```

```
##       gcatP
## Year   preferred+ preferred-
##   2013         16         82
##   2014         18        122
##   2015         15         52
##   2016         11         96
##   2017          9         26
```

```
chisq.test(lmb.LFP)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lmb.LFP
## X-squared = 8.2649, df = 4, p-value = 0.08234
```

```
(ps.P <- c(chisq.test(lmb.LFP[c(1,2),])$p.value,  ### 2013-2014
           chisq.test(lmb.LFP[c(1,3),])$p.value,  ### 2013-2015
           chisq.test(lmb.LFP[c(1,4),])$p.value,  ### 2013-2016
           chisq.test(lmb.LFP[c(1,5),])$p.value,  ### 2013-2017
           chisq.test(lmb.LFP[c(2,3),])$p.value,  ### 2014-2015
           chisq.test(lmb.LFP[c(2,4),])$p.value,  ### 2014-2016
           chisq.test(lmb.LFP[c(2,5),])$p.value,  ### 2014-2017
           chisq.test(lmb.LFP[c(3,4),])$p.value,  ### 2015-2016
           chisq.test(lmb.LFP[c(3,5),])$p.value,  ### 2015-2017
           chisq.test(lmb.LFP[c(4,5),])$p.value)) ### 2016-2017
```

```
## Warning in chisq.test(lmb.LFP[c(4, 5), ]): Chi-squared approximation may be
## incorrect
```

```
(p.val.P <- p.adjust(ps.P))
```

```
##  [1] "13-14" "13-15" "13-16" "13-17" "14-15" "14-16" "14-17" "15-16"
##  [9] "15-17" "16-17"
```

```
##     Year p-value Adjusted p
## 1  13-14  0.5724     1.0000
## 2  13-15  0.4378     1.0000
## 3  13-16  0.2837     1.0000
## 4  13-17  0.3329     1.0000
## 5  14-15  0.1212     0.8485
## 6  14-16  0.6716     1.0000
## 7  14-17  0.1048     0.8387
## 8  15-16  0.0498     0.4565
## 9  15-17  0.8964     1.0000
```

```
## 10 16-17  0.0456      0.4565
```

The PSD-P of largemouth bass is not different for any years during 2013 - 2017 (Chi-Squared, $X^2 = 8.26$, df = 4, p = 0.08). The adjusted p-values show no difference in the PSD-P between years (2013 - 2017). T